buildings

MDPI

Article

Reimagining Residential Planning: A New Paradigm Using Grasshopper Combined with Fine-Tuned Diffusion Models

Kai Zhang [1], Xu Zhang [2,*] and Junming Chen [3,*]

[1] College of Animation, Jinling Institute of Technology, Nanjing 211169, China; zhangkai@jit.edu.cn
[2] Faculty of Innovation and Design, City University of Macau, Macau 999078, China
[3] Faculty of Humanities and Arts, Macau University of Science and Technology, Macau 999078, China
* Correspondence: u24092110508@cityu.edu.mo (X.Z.); jmchen@must.edu.mo (J.C.)

Abstract

Rapid global urbanization has increased the demand for residential construction, rendering traditional design methods insufficient. This study proposes an innovative approach integrating parametric techniques with fine-tuning diffusion models to automate the modeling and rendering of residential area planning. Specifically, this method first utilizes Grasshopper programming to develop a parametric design system that can automatically create 3D models of residential areas. Subsequently, a new dataset was constructed to fine-tune the diffusion model, enabling it to generate high-quality renderings that conform to specific architectural styles, types, and viewpoints. The proposed method transforms the traditional scheme refinement, modeling, and rendering processes into a generative design process. In a subjective evaluation of 2880 generated images by ten designers, the proposed method obtained higher scores than Midjourney, DALL E3, and Stable Diffusion in usability (81%), building style consistency (87%), and realism (92%). Results indicate that this method can efficiently generate diverse design options at the preliminary design stage. This study provides a practical workflow for the rapid generation of residential area models and renderings.

Keywords: Grasshopper; generative design; fine-tuning diffusion model; residential area planning; workflow improvement

1. Introduction

1.1. Background and Motivation

Research indicates a mutually reinforcing relationship between urbanization and industrial transformation. Industrial transformation and expansion create considerable job opportunities and increase urban population density, accelerating urbanization [1–3]. Currently, most countries are urbanizing, and housing new urban residents has become urgent.

However, traditional residential area planning and design face challenges in the current urbanization process, with low design efficiency and homogeneous architectural styles [4–6]. Firstly, conventional residential area planning and design adopt a linear workflow for 3D modeling and rendering, which is complex and difficult to modify, resulting in low design efficiency [7]. Secondly, individual designers with limited design styles tend to design many residential areas with similar architectural styles, and the architectural appearance of different urban areas lacks diversity [8,9]. The accelerating urbanization urgently requires new technologies to innovate traditional residential area planning methods for improved design efficiency and enriched architectural design styles.

Check for updates

Academic Editor: Chi-Ming Lai

Received: 15 September 2026
Revised: 6 October 2026
Accepted: 12 October 2026
Published: 13 October 2026

Copyright: © 2026 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license.

1.2. Research Gap

The development of parametric technology and AI has provided new possibilities for approaching residential area design issues. Existing research has explored the potential of parametric technology [10–12] and the feasibility of AI-based image generation technology in designing [13–16]. However, applying these technologies in residential area planning and design still faces certain limitations. Firstly, residential area planning is a complex system engineering task requiring attention to the entire design process, while previous parametric studies often focused on optimizing specific design nodes [17–19]. Additionally, existing research attempts to employ pre-trained diffusion models to generate design renderings, but the results are unsatisfactory [20,21]. This is because traditional diffusion models are trained on large quantities of text-image pairs from the internet, lacking annotations with specialized domain knowledge [22–25]. In particular, little research has applied fine-tuned diffusion models to residential area planning and design. A single programming technology or an AI alone can not surmount the complex challenges in residential area planning.

1.3. Problem Statement and Objectives

The planning and design of a residential area refers to the rational layout of residential buildings, supporting facilities, and public spaces on a specific piece of land, considering social, economic, and environmental factors. From the designer's perspective, this process has two main challenges. First, current residential area planning primarily relies on manual completion by designers, including layout planning, three-dimensional modeling, and effect rendering. If a single designer were to undertake these processes independently, it would take more than a month, highlighting the urgent need to enhance design efficiency. Second, due to development cycle constraints and limitations in designer experience, it is frequently impractical to provide multiple design alternatives for a single project. This limitation restricts both the quality and diversity of design outcomes. Given the challenges associated with efficiency and diversification of design styles, there is an urgent need for a new approach to generate residential area planning and design. Specifically, this approach should facilitate the bulk generation of three-dimensional models and renderings, thereby supporting initial design iterations and finalizing design schemes.

This study proposes a new residential area design method based on parametric technology and fine-tuned diffusion models. It integrates the advantages of parametric programming and AI to automatically generate 3D residential area models and renderings in specific architectural styles according to the inputted requirements. This process is intended to reduce the time spent on repetitive drawing, modeling, and rendering at the preliminary design stage, thereby improving the efficiency of generating residential area design alternatives.

Fine-tuning diffusion models is essential because conventional diffusion models, although trained on extensive internet data, lack specialized annotations in specific domains. In specialized fields, conventional diffusion models suffer from a scarcity of training samples, leading to suboptimal performance. Fine-tuning diffusion models can effectively introduce new knowledge, enhancing their performance and understanding of specialized knowledge at a lower cost. Consequently, this study adopts the method of fine-tuning diffusion models to generate higher-quality planning renderings for residential areas.

1.4. Methodology Overview

The proposed method combines parametric design with an image generation model to directly generate 3D residential area models and renderings, namely, generating residential designs from parameters and text prompts (GRDPT). It supports rapid design regeneration and modification and incorporates algorithm design and user application stages.

The algorithm design phase includes the parametric design and image generation modules. In the parametric design module, three-dimensional models of residential areas are automatically generated based on input parameters. Researchers encode the logic of residential area planning design using Grasshopper, incorporating adjustable parameters to accommodate different design requirements. Next, the image generation module converts the three-dimensional models into renderings. This module involves constructing a dataset of residential area planning and fine-tuning the diffusion model. First, we have assembled the Residential Area Planning Dataset (RAPD-36), which comprises 1326 images of residential areas. These maps encompass a variety of styles, building types, and viewpoints. Subsequently, a multi-concept fine-tuning strategy based on the prior-preservation loss of DreamBooth [26] is adopted, in which architectural styles, building types, and viewpoints are encoded in the text prompts. By employing this strategy and RAPD-36, a multi-style architectural residential planning rendering model (MARPR) is trained. The fine-tuned MARPR is intended to alleviate the shortcomings of traditional models in professional design by generating batch renderings of residential areas in specified styles, types, and viewpoints.

In the application phase, users input design requirements as parameters into the Grasshopper program, which rapidly generates three-dimensional models of residential area planning within seconds. Real-time adjustments can be made to modify these models as needed. Subsequently, users can export images from selected angles of the three-dimensional models and provide textual descriptions of design requirements to the MARPR model, generating renderings in the desired style. Through this process, designers can obtain the three-dimensional model of the residential area and its corresponding rendering.

The proposed method reduces the steps required in traditional methods for creative design, drawing, modeling, and rendering, thereby allowing designers to efficiently generate residential area planning. It is expected to improve design efficiency and enrich the diversity of residential area designs, further promoting the intelligent development of the design industry. The research framework is detailed in Figure 1.

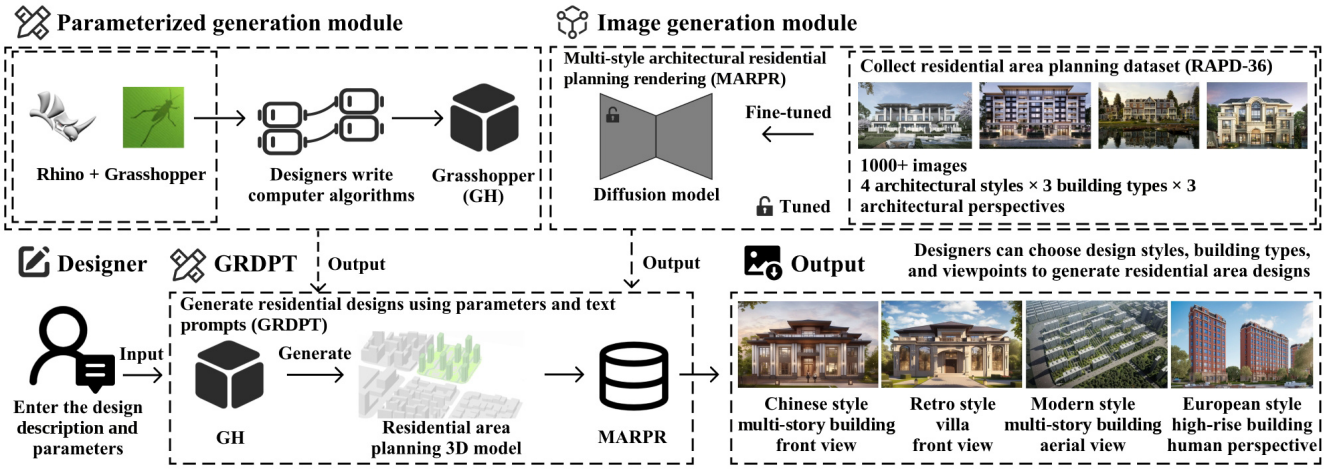

Figure 1. Research framework. The designed approach allows the parametric generation of residential area plan models and employs a fine-tuned diffusion model to produce residential area plans. It only requires design requirements to obtain 3D models and renderings of residential area designs in different building styles.

The advantage of the proposed GRDPT method lies in the automatic generation of 3D residential area models through parametric design and the production of renderings in specified design styles through MARPR. GRDPT makes it possible to quickly and controllably produce 3D residential area models and the corresponding renderings. Figure 2

shows the renderings of the controllable residential area designs generated by GRDPT according to control conditions such as building viewpoint, building style, building type, and segmented image.



Figure 2. GRDPT generates designs and corresponding renderings. It can generate 3D models and renderings of a residential area based on specific prompts and segmented images. (**A**) Retro-style and European-style residential buildings shown from the human perspective and the front view; (**B**) High-rise and multi-story residential areas shown from the aerial view and the front view; (**C**) Modern-style and Chinese-style designs for multi-story buildings and villas; (**D**) Retro-style and modern-style high-rise renderings generated under segmentation control.

In the evaluation by ten designers, the proposed approach obtained higher scores than the compared image generation models in usability, building style consistency, and realism. Through a controllable process, the proposed approach automates the generation of 3D residential area models and their renderings at the preliminary design stage, while design decisions remain with the designers. By combining parametric generation and fine-tuned diffusion models, this study provides a practical workflow for applying AI-assisted visualization to residential area planning.

Regarding the comparison of methodological differences, several methods have been proposed to adapt pre-trained diffusion models to specific tasks. DreamBooth [26] fine-tunes a diffusion model on a small set of images and uses a prior-preservation loss to retain the knowledge of the original model, but it is usually applied to a single subject or concept. LoRA [27] freezes the pre-trained weights and trains low-rank matrices inserted into the network, which greatly reduces the number of trainable parameters, and is generally used to learn a single style or subject. Custom Diffusion [28] enables several concepts to be learned jointly, but each concept is typically a specific object or style learned from a few images. ControlNet [29] adds spatial conditions, such as edge or segmentation maps, to control the composition of the generated image, but it does not learn new styles. In architectural design, existing Grasshopper-based workflows mainly automate modeling and

performance optimization, while material assignment and rendering are still performed separately. In contrast, GRDPT applies fine-tuning to the joint learning of three compositional attributes, namely architectural style, building type, and viewpoint, in a single model, and uses the segmentation images exported from Grasshopper as the spatial condition, forming a complete workflow from parametric modeling to rendering.

In addition, our team has conducted a series of studies on AI-assisted architectural and interior design generation. Fine-tuned diffusion models were used to generate architectural designs and interior designs from text descriptions [8,13]. Aesthetic evaluation was further incorporated into diffusion-based interior design generation to make the generated designs more usable [14]. An improved control network was developed to generate interior designs that match the indoor structure [30], and a trained model was used to generate interior design videos of various styles from texture-free 3D models [31]. In addition, Grasshopper was used to achieve automated interior design and manufacturing [10]. Each of these studies addressed a specific aspect of design generation. Different from them, the present study combines Grasshopper-based parametric modeling with diffusion model fine-tuning to reshape the design workflow of residential area planning, covering layout generation, 3D modeling, and rendering.

1.5. Main Contribution

The main contributions of this study are as follows:

1. A design workflow integrating Grasshopper-based parametric modeling with a fine-tuned diffusion model has been established to accelerate the exploration of residential area schemes.
2. A rule-based Grasshopper program has been developed to generate 3D residential area models, whose semantic color layers serve as control images for image generation.
3. RAPD-36, a dataset of 1326 residential area images annotated by architectural style, building type, and viewpoint, has been constructed.
4. Fine-tuning has been extended from single-subject learning to the joint learning of architectural style, building type, and viewpoint in a single, locally deployable model.
5. The usability and limitations of the proposed workflow have been examined through comparative experiments and a subjective designer evaluation.

2. Related Work

2.1. Residential Area Planning

Residential areas are typically designated urban areas for living and residential facilities, public green spaces, and roads. Constructing residential areas instead of standalone buildings can efficiently provide citizens with housing and living facilities, thus meeting the needs of a rapidly developing city. In developing countries with rapid urbanization, constructing residential areas can accommodate new urban residents, thereby promoting national economic prosperity [32–34]. Therefore, advancing the construction of high-quality residential areas is crucial.

However, traditional residential area design methods have limited efficiency and creativity. Regarding design efficiency, conventional residential area design processes include design requirement analysis, plan layout drawing, design style selection, 3D model construction, material texture selection, and final rendering. Each design phase requires a significant amount of time from designers [35,36]. Additionally, residential area planning and design requirements frequently change, and designers spend considerable time on repeated modifications [7,30,31]. In terms of creative design, each designer excels in a limited range of styles, and mastering all design styles is unrealistic. Meanwhile, retraining costs are very high. Therefore, it is demanding for designers to create diverse architectural

styles [8,37,38]. In summary, design efficiency and creative design issues are the main challenges in residential area planning [4–6,39]. Thus, addressing these challenges with more efficient and intelligent design methods has become imperative.

2.2. Parametric Design

Parametric design relies on parameters and variables to control design outcomes. Designers can quickly generate different design schemes by adjusting the parameters [10,40,41]. It is widely used in various fields, including product design [42], architectural design [43], and artistic creation [44]. In architectural design, parametric modeling and design are often implemented using Rhino and Grasshopper. Compared with traditional methods, parametric technology allows designers to generate and adjust design schemes almost in real-time. By altering the parameters, designers can quickly evaluate and modify designs, thereby enhancing work efficiency [40,45–47].

Parametric design based on Grasshopper mainly involves components, sliders, and connections. Components are collections of preset computer instructions, sliders adjust the parameters, and connections link different components to form more complex operational instructions. Figure 3 shows a standard parametric design process. First, the designer selects the required components. Then, sliders and text input boxes are added as needed to reserve adjustable parameters. Finally, all the components and sliders are linked through connections to establish an automated workflow. This design approach allows designers to update the design in real-time by adjusting the sliders and text inputs, flexibly adjusting and optimizing the design scheme, thereby accelerating design iteration and optimization [48]. Additionally, it provides an interactive visual design environment for designers to easily create complex design algorithms by simply dragging and connecting components, thereby facilitating design generation through programming [49]. Considering the above advantages, Grasshopper programming is adopted to program the automated generation of 3D residential area models.

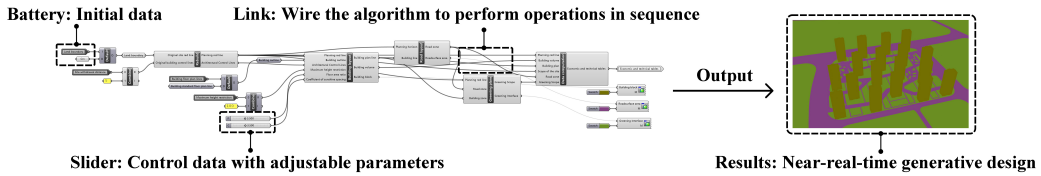

Figure 3. Schematic diagram of Grasshopper-based parametric design process.

2.3. Fine-Tuned Diffusion Model

Although parametric design methods can achieve automated modeling in residential area planning and design, material mapping and rendering are also indispensable in the design process. Traditional design methods usually assign material maps to 3D models and render them to produce design images. The drawbacks are the inefficient identification of suitable maps and rendering and the difficulties for modification. Therefore, this study proposes surmounting this issue through image generation based on fine-tuned diffusion models.

Mainstream image generation models include Stable Diffusion [50], Midjourney [51], and DALL E3 [52]. These models are trained on a large amount of text and image data from the internet, enabling the generation of images from text prompts. As an open-source model, Stable Diffusion allows free customization and improvement by developers, artists, and researchers worldwide according to their needs, thus promoting its application and development in different fields [31,53].

Despite their broad application prospects in design, pre-trained Stable Diffusion models may perform poorly on specialized tasks because their training data lack sufficient

domain-specific annotations. In residential area planning, generated images may not consistently follow the specified architectural styles, building types, and viewpoints. Fine-tuning offers a way to improve adherence to these requirements by adapting the pre-trained model using domain-specific images and corresponding text descriptions. In this study, we retain the original DreamBooth loss function and network architecture and fine-tune the model on RAPD-36 to support residential area rendering generation.

2.4. The Role of Datasets in AI

Datasets are crucial in AI development [54,55]. For example, the ImageNet dataset contains over 14 million high-definition images, covering over 20,000 categories. ImageNet has enabled studies and comparisons based on the same dataset, thereby significantly advancing the field of image vision [56]. Another example is the large-scale source image and text datasets proposed by LAION, such as LAION-400M and LAION-5B. The massive and diverse data of LAION-5B make it an ideal choice for training high-performance visual-text pairing models, effectively promoting research on image generation from text [57]. These large datasets provide rich resources for training and evaluating new models, driving progress and innovation in AI.

The above case demonstrates the importance of datasets for developing related fields. However, in residential area planning, the lack of publicly available datasets with diverse design styles and building type annotations has led to slow progress in research on residential area planning image generation. This study aims to construct a new residential area planning dataset to address the issue of dataset scarcity, thereby promoting the application and development of the Stable Diffusion (SD) [50] generation model in residential area planning. This initiative will provide richer data resources for research and practice in residential area planning and is expected to drive progress in related technologies and methods.

3. Methodology

This study proposes a novel GRDPT method to address the inefficient planning and design processes [58] and the creative design challenges [9]. GRDPT includes a parametric 3D model generation module and a fine-tuned MARPR model. The following sections detail the proposed method. Section 3.1 constructs the parametric design program through Grasshopper programming to allow designers to directly generate residential area planning schemes by modifying preset parameters. Section 3.2 builds the RAPD-36 dataset containing architectural styles. Section 3.3 presents the multi-concept fine-tuning method for the model using the RAPD-36 dataset. Finally, Section 3.4 demonstrates the complete workflow of the proposed method. To compare the performance of different AI-generated image tools, this study used Midjourney 5.2, DALL-E 3, and Stable Diffusion 1.5 to generate Figures for a visual comparison of model performance.

3.1. Grasshopper-Based Parametric Design Program Implementation

Through Grasshopper programming, the generative residential area planning program is implemented in six steps. The first step involves sketch reconstruction, where the Grasshopper algorithm repairs issues such as duplicate segments and broken lines in CAD drawings, yielding smooth and closed curves. The second step involves the automated generation of the residential area layout. Specifically, the Grasshopper algorithm arranges pre-selected building floor plans along the edges of the residential area, trimming areas unsuitable for construction, and places buildings in the remaining space row by row until no further building can be placed. Generally, the program's automatic layout proceeds from south to north and follows the sunlight spacing, building interval, and setback

constraints specified by the input parameters. This step is based on preset programming logic rather than artificial intelligence algorithms; therefore, the initial layout satisfies only the constraints encoded as parameters, and other regulatory requirements, such as fire safety, still need to be checked by designers. The third step involves importing our pre-written GH program to generate a 3D model of the buildings. This process involves recognizing the wall and window lines drawn on different layers of the selected building layout diagram. The generated walls and windows are solid volumes, incorporating only essential building elements. This approach is informed by preliminary experimental tests in this study, which revealed that excessive structure would limit the diversity of building generation in later stages. Furthermore, overly complex structures can slow down computer performance when displaying the 3D model of the entire community plan. The fourth step generates the internal road network of the residential area by identifying building locations, extending corresponding roads, and connecting them. The fifth step generates greenery, defining the unused plots into public spaces and landscaped green areas. The final step outputs the generated elements in groups according to preset semantic color layers, thus completing the 3D residential area planning. Figure 4 shows the full logical flow of the Grasshopper-based generative residential area planning program.

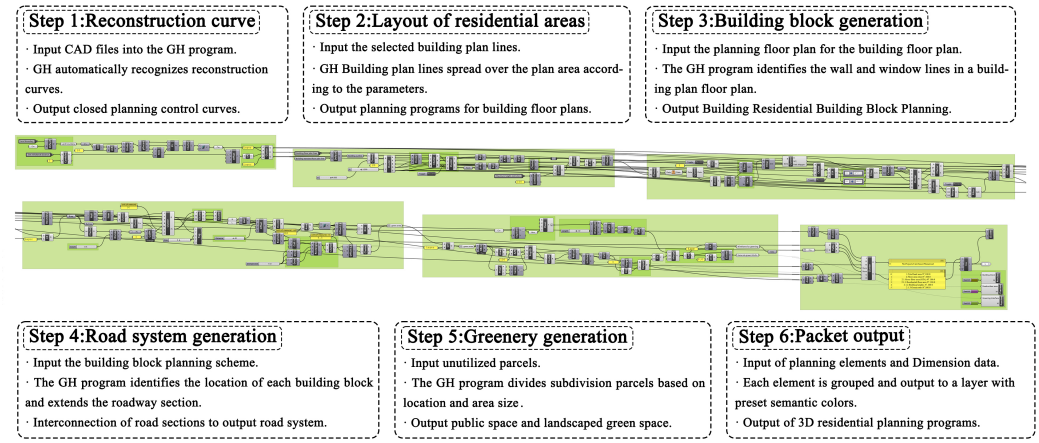

Figure 4. Grasshopper-based parametric design program logic.

The Grasshopper program requires necessary and optional parameters to generate 3D residential area designs, including floor area ratio, coefficient of sunshine spacing, maximum height restriction, and building interval. The optional parameters are preset with default values but can be modified by the designer. These optional parameters include building story height, roadway width, sidewalk width, building setback angle, and site withdrawal distance. All parameters are listed in Table 1, together with their ranges, steps, and default values. The default values are presets for a typical residential area design and can be adjusted by the designer through the sliders according to the requirements of the specific project.

The program does not automatically check whether a parameter combination is valid. When the parameters are invalid or conflict with one another, for example, when the spacing and withdrawal requirements leave no room to place a building on the site, the Grasshopper program cannot generate the 3D model. Such cases are identified by the designer, who then adjusts the corresponding sliders until a valid 3D model is generated. Because the model is regenerated within seconds after each adjustment, this manual correction can be completed quickly.

Table 1. Adjustable parameters of the pre-written Grasshopper-based program and their default values.

	Parameters	Range	Steps	Default
X_1	Floor area ratio	$[0, 10]$	0.1	2.0
X_2	Coefficient of sunshine spacing	$[0, 2]$	0.1	1.2
X_3	Maximum height restriction	$[0, 100]$	0.1 m	54.0 m
X_4	Building interval	$[0, 20]$	0.1 m	13.0 m
X_5	Building story height	$[0, 6]$	0.1 m	3.0 m
X_6	Roadway width	$[0, 12]$	0.1 m	6.0 m
X_7	Sidewalk width	$[0, 6]$	0.1 m	1.5 m
X_8	Building setback angle	$[0, 360]$	1°	0°
X_9	Site withdrawal distance	$[4, 14]$	0.1 m	5.0 m

3.2. RAPD-36 Dataset Construction

To fulfill the need for residential area planning data, professional designers were invited to collect residential area planning images that are publicly available on the internet, mainly retrieved through Google image search and from the gooood design website [59]. Only images that met the following criteria were collected: (1) the architectural style, building type, and viewpoint of the image can all be clearly identified, so that the image corresponds to exactly one of the 36 predefined categories; (2) both the width and the height of the image are greater than 512 pixels; and (3) the image contains no watermark. Duplicate images were also removed during collection. In total, 1694 images were collected. RAPD-36 includes residential area design images from four architectural styles: Chinese style, retro style, European style, and modern style, reflecting mainstream residential area planning designs on the market. Additionally, each image was classified as high-rise, multi-story building, or villa based on building type. Furthermore, the images in the RAPD-36 dataset were divided into aerial view, front view, and human perspective based on architectural viewpoint. In total, the dataset has 36 categories.

The copyright of the collected images belongs to their original owners. Therefore, the images are used only locally for non-commercial academic research and are not redistributed, and the trained model is also used only for research purposes. This practice is consistent with the common way in which image datasets are constructed in computer vision research. For example, ImageNet [56] was built by collecting candidate images from the internet through several image search engines, removing duplicate images, and verifying the labels through multiple human annotators. Because the ImageNet team does not own the copyright of these images, access to them is provided only for non-commercial research and educational purposes [60]. ImageNet has since become one of the most widely used datasets for training and evaluating image recognition models and has greatly advanced the development of computer vision [56]. Similarly, the LAION-5B dataset [57], which was used to train Stable Diffusion [50], the base model of this study, was also built from publicly available web images. RAPD-36 follows the same principles on a smaller scale: the images were collected from public web sources, screened and verified by multiple designers, and used only for non-commercial research without redistribution.

Before labeling the RAPD-36 dataset, we trained all invited image annotators despite their knowledge of common architectural styles. The annotators were three designers: two postgraduate students in design with 7 and 8 years of study in the field, respectively, and one design faculty member with 15 years of professional experience. Specifically, we required all designers to collaboratively label 100 images of different architectural styles, identifying any inconsistencies in the labeling categories. Subsequently, the designers discussed the labeling methods for these images and reached a consensus on the following labeling rules. Chinese style is characterized by sloped roofs and traditional Chinese decorative elements; retro style by stone or stone-like walls and classical compositions;

European style by red bricks, dormer windows, and decorative details; and modern style by simple volumes and large areas of glass. Villas are low-rise detached or terraced houses, multi-story buildings have 4–6 stories, and high-rise buildings have more than 6 stories. Aerial views are taken from above the roofs, front views face the main facade, and human perspectives are eye-level views from within the site. In the subsequent labeling process, each image was labeled independently by the three designers in terms of architectural style, building type, and viewpoint. Only images that received identical labels from all three designers were retained, and images with inconsistent labels were removed. Of the 1694 collected images, 1326 (78.3%) received identical labels from all three designers and were retained to form the RAPD-36 dataset. Most disagreements concerned architectural style, because some buildings combine features of several styles.

Table 2 shows the number of images in each of the 36 categories of the RAPD-36 dataset. In terms of design style, there are 493 modern-style, 319 retro-style, 292 Chinese-style, and 222 European-style images. Regarding building type, there are 563 high-rise buildings, 399 villas, and 364 multi-story buildings. In terms of architectural viewpoint, there are 493 aerial-view, 414 front-view, and 419 human-perspective images. The number of images per category ranges from 13 (retro-style multi-story buildings in front view) to 90 (modern-style high-rise buildings in aerial view), which reflects the different availability of such images on the internet. No re-balancing was applied during training, for two reasons. First, previous studies have shown that a few images are sufficient for fine-tuning a diffusion model to learn a new concept; for example, DreamBooth learns a specific subject from only 3–5 images [26], whereas even the smallest category in RAPD-36 contains 13 images. Second, the three attributes are shared across categories. For example, the 13 retro-style multi-story front-view images share the same style with all 319 retro-style images, the same building type with all 364 multi-story images, and the same viewpoint with all 414 front-view images. Because each attribute is learned through the corresponding words in the text prompts, the model can learn each attribute from all images that contain it, so that categories with fewer images can benefit from the knowledge learned from other categories. As a result, each style, building type, and viewpoint is learned from at least 222 images. In addition, the same number of generated images (20) was used for each of the 36 categories in the evaluation (Section 4.5). Figure 5 shows the high-quality and low-quality image samples of the RAPD-36 dataset.



Figure 5. Sample images from the RAPD-36 dataset.

Table 2. The number of images in different categories of RAPD-36.

Viewpoint	High-Rise			Multi-Story Building			Villa		
	Av [1]	Fv [1]	Hp [1]	Av	Fv	Hp	Av	Fv	Hp
Chinese style	51	49	31	36	25	19	29	20	32
Retro style	61	39	30	35	13	55	38	27	21
European style	35	28	28	23	19	23	17	19	30
Modern style	90	53	68	44	44	28	34	78	54

[1] Av: Aerial view; Fv: Front view; Hp: Human perspective.

3.3. Diffusion Model Fine-Tuning

This study adopts a multi-concept fine-tuning strategy that jointly learns architectural style, building type, and viewpoint. With this strategy and the RAPD-36 dataset, the fine-tuned MARPR model can effectively generate batches of residential area planning images that meet specific architectural style, type, and viewpoint demands. Specifically, residential area planning images from RAPD-36 are input into the diffusion model together with their text prompts for training. As a result, the model can grasp critical information related to architectural style, type, and viewpoint during training and generate residential area planning designs that meet specific requirements. This method is intended to enhance the practicality of diffusion models in architectural design, enabling designers to input textual design requirements for automatic design image generation.

The fine-tuning objective in Equation (2) follows the DreamBooth framework with prior preservation [26]. It combines the denoising reconstruction objective in Equation (1) with a prior-preservation term:

$$\mathbb{L}_{\boldsymbol{Y},\boldsymbol{g},\boldsymbol{\epsilon},\boldsymbol{t}}[w_t||\hat{Y}_\theta(\alpha_t Y + \sigma_t \epsilon, \boldsymbol{g}) - Y||_2^2] \tag{1}$$

where $\mathbb{L}$ denotes the expected loss over training images, text conditions, sampled noise, and diffusion timesteps. Y is the clean image representation, ϵ is Gaussian noise, and α_t and σ_t determine the noise level at timestep t. The trainable denoising model $\hat{Y}_\theta$ predicts the clean representation from the noisy input $\alpha_t Y + \sigma_t \epsilon$, conditioned on the text vector $\boldsymbol{g}$. Training updates the model parameters θ to minimize the reconstruction error. The timestep weight w_t determines the contribution of errors at different noise levels to the loss.

The fine-tuning loss function is expressed in (2):

$$\begin{aligned} &\mathbb{L}_{\boldsymbol{Y},\boldsymbol{g},\boldsymbol{\epsilon},\boldsymbol{\epsilon}',\boldsymbol{t}}[w_t||\hat{Y}_\theta(\alpha_t Y + \sigma_t \epsilon, \boldsymbol{g}) - Y||_2^2 \\ &+\lambda w_{t'}||\hat{Y}_\theta(\alpha_{t'} Y_{pr} + \sigma_{t'} \epsilon', \boldsymbol{h_{pr}}) - Y_{pr}||_2^2] \end{aligned} \tag{2}$$

Compared with Equation (1), Equation (2) incorporates building style, building type, and building viewpoint through the text vector $\boldsymbol{g}$ and introduces the prior-preservation loss of DreamBooth [26] to enhance control over the generation results of the diffusion model. Equation (2) consists of two parts. The first part measures the difference between the generated image and the actual image. $\hat{Y}_\theta$ represents the new diffusion model, incorporating building style, type, and viewpoint through the text vector $\boldsymbol{g}$. The difference between the generated image and the actual image Y leads to the first loss component. The second part is the prior-preservation loss, which compares the image generated by the new diffusion model (i.e., $\hat{Y}_\theta(\alpha_{t'} Y_{pr} + \sigma_{t'} \epsilon', \boldsymbol{h_{pr}})$) with the image generated by the pre-trained diffusion model (i.e., Y_{pr}). With a smaller difference between the two, the newly trained model better retains the foundational knowledge of the original model. λ is a fixed weight used to adjust the weights of these two parts of the loss function for better generation results. Combining the first and second loss functions enables the new diffusion model to learn

the building style, type, and viewpoint while maintaining the knowledge of the original diffusion model. As a result, the fine-tuned MARPR model can generate controllable residential area planning designs in terms of building style, type, and viewpoint.

It should be clarified that Equation (2) is the original DreamBooth loss [26]; no modification was made to the loss function or the network architecture. The difference from DreamBooth lies only in how it is used: instead of binding one subject learned from 3–5 images to a single identifier, all RAPD-36 images are trained in one run with structured prompts composed of style, building-type, and viewpoint words, so that the 36 combinations are learned jointly in a single model. Therefore, this study does not claim an algorithmic contribution; its contribution lies mainly in the workflow that links Grasshopper-based parametric modeling with the fine-tuned model, and in examining the usability and limitations of this workflow.

It should be noted that architectural style, building type, and viewpoint are not implemented as three independent loss terms. Following a multi-concept training strategy, each training image Y in RAPD-36 is paired with a structured text prompt $p(s, b, v)$, where s is one of the four architectural styles (Chinese, retro, European, and modern), b is one of the three building types (high-rise, multi-story, and villa), and v is one of the three viewpoints (aerial view, front view, and human perspective), e.g., "Chinese-style residential building, multi-story building, human perspective". The text vector is obtained as $\boldsymbol{g} = \tau(p(s, b, v))$, where τ denotes the CLIP text encoder of Stable Diffusion. Therefore, the losses associated with the three attributes are contained in the first term of Equation (2): for each image, the reconstruction error is minimized under the condition of its own style, type, and viewpoint words, so that the model learns the visual features corresponding to each attribute, and all 36 combinations are learned jointly in a single model.

In Equation (2), ϵ and ϵ' are Gaussian noise sampled from a standard normal distribution, t and t' are randomly sampled diffusion timesteps, and α_t and σ_t are the noise schedule coefficients that determine the noise level at timestep t. Y_{pr} denotes the class images generated by the pre-trained Stable Diffusion 1.5 from the class prompt $\boldsymbol{h_{pr}}$, and θ denotes the trainable parameters of the diffusion model. w_t and $w_{t'}$ are timestep weights, and λ is a fixed weight that balances the reconstruction loss and the prior-preservation loss.

The training procedure is as follows. First, 400 class images Y_{pr} were generated by the pre-trained Stable Diffusion 1.5 using the class prompt $\boldsymbol{h_{pr}}$. Second, each image in RAPD-36 was paired with its structured prompt $p(s, b, v)$ according to its annotation. Third, during training, batches of RAPD-36 images and class images were sampled, noise was added to them at randomly sampled timesteps, and the loss in Equation (2) was computed with $\lambda = 1$. The model parameters θ were then updated with the AdamW optimizer, which is the default optimizer of the training script. This process was repeated until each RAPD-36 image had been used 100 times, and the model obtained at the end of training was used as MARPR.

The resources for MARPR training are detailed as follows. The RAPD-36 dataset and the loss function in Equation (2) were used to train MARPR. The training was conducted on a Windows 10 PC with 64 GB of RAM and an NVIDIA 3090 GPU. PyTorch 2.9.0 was used for model training, with each image undergoing 100 iterations. The training infrastructure was Stable Diffusion 1.5, and the image preprocessing method involved scaling the input images to a maximum dimension of 512 pixels. Data augmentation was performed through mirroring. The learning rate for the model was set at 0.000001, with a batch size of 24. Xformers and FP16 were utilized for accelerated computation. The training took a total of 8 h.

3.4. Design Generation and Modification

This section compares GRDPT with the traditional residential area design workflow. First, the traditional workflow must conceive a design plan and draw a layout of the residential area. Then, a corresponding 3D model must be built, and material textures must be selected for design rendering. This design process is linear and cumbersome. Meanwhile, modifying the design plan requires re-executing the entire design process, hindering design efficiency and quality.

In contrast, GRDPT only requires input of site drawings, parameters, and textual descriptions to generate 3D models and renderings of the residential area. Specifically, designers initially combine layered CAD drawings with pre-written Grasshopper (GH) programs to generate a 3D model of the residential area. Subsequently, they select specific views for rendering through GH's visual interface. Designers can also install plugins that capture real-time screen images, enabling them to obtain live images and automate their input into MARPR. This method allows users to complete the modification of the 3D model and the generation of the rendered images within a single GH interface. The entire rendering process can be completed in just a few seconds. If users are unsatisfied with the results, they can adjust parameters in GH, replace the floor plan to regenerate a new residential area layout and 3D model, or create new planning renderings by modifying the textual prompts. GRDPT can reduce the time that designers spend on drawing and modeling and provide them with various candidate plans more efficiently. Specifically, GRDPT reduces the repetitive drawing, modeling, and rendering work in traditional design methods, while the conceptual design decisions remain with the designers. Figure 6 visually compares the two design methods.

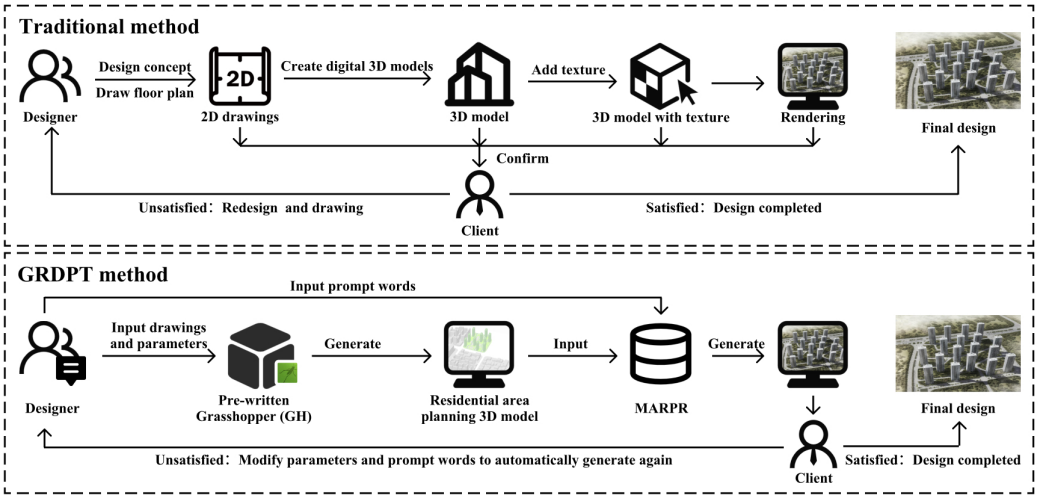

Figure 6. Comparison between the traditional design method and GRDPT.

4. Experiments and Results

This section presents the designs generated by GRDPT and the corresponding evaluation results, including the 3D models and renderings. Section 4.1 demonstrates the generation of 3D residential area models based on Grasshopper. Section 4.2 presents the residential area renderings generated by MARPR. Section 4.3 showcases the metrics proposed by industry experts for evaluating the quality of the generated images. Section 4.4 shows the visual evaluation results. Section 4.5 presents the quantitative evaluation results. Section 4.6 details the design specifics generated by GRDPT.

4.1. Generating Residential Area Plans Based on Grasshopper

A Grasshopper program is written to generate residential area plans. With this program, designers only need to input a CAD site file and set the necessary parameters to generate the corresponding 3D residential area model. Subsequently, designers can freely switch perspectives to observe the generated design and choose the appropriate view angle to output the 3D model as a semantic segmentation image. Figure 7 shows the process for the Grasshopper program to generate a residential area plan. First, the designer inputs the CAD site file. Second, the designer inputs the necessary generation parameters into the Grasshopper program, including the floor area ratio, coefficient of sunshine spacing, maximum height restriction, building interval, and other optional parameters. Then, the Grasshopper program automatically calculates and generates the building layout based on the input parameters. Next, the Grasshopper program generates driveways and sidewalks based on the building layout. Subsequently, the Grasshopper program generates greenery based on the remaining space. Finally, a 3D model with semantic segmentation is generated.

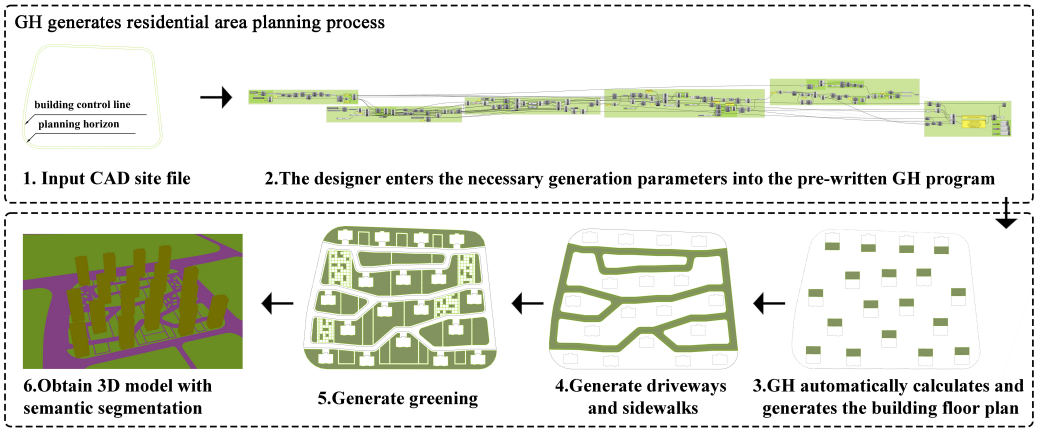

Figure 7. Grasshopper-based residential area design generation process.

4.2. Generating Designs with Different Styles, Types, and Perspectives

Figure 8 demonstrates the performance of GRDPT in generating residential area renderings with different architectural styles, types, and perspectives. The X-axis represents the combination of three residential building types and three architectural perspectives, while the Y-axis corresponds to four residential architectural styles, totaling 36 combinations. GRDPT can generate residential area renderings based on the specified architectural style, type, and perspective, thereby effectively controlling the generated results. The GRDPT-generated architecture styles differ significantly. For example, the generated Chinese-style buildings feature sloped roofs, the retro-style buildings often have stone walls, and the European-style buildings have red bricks, dormer windows, and decorations. In contrast, the modern-style buildings feature large areas of glass. Additionally, the generated designs include rich design details. The quality of these renderings is close to actual usage needs. Thus, designers can enjoy accelerated design and decision-making processes based on the generated 3D residential area models and renderings (Figure 8).

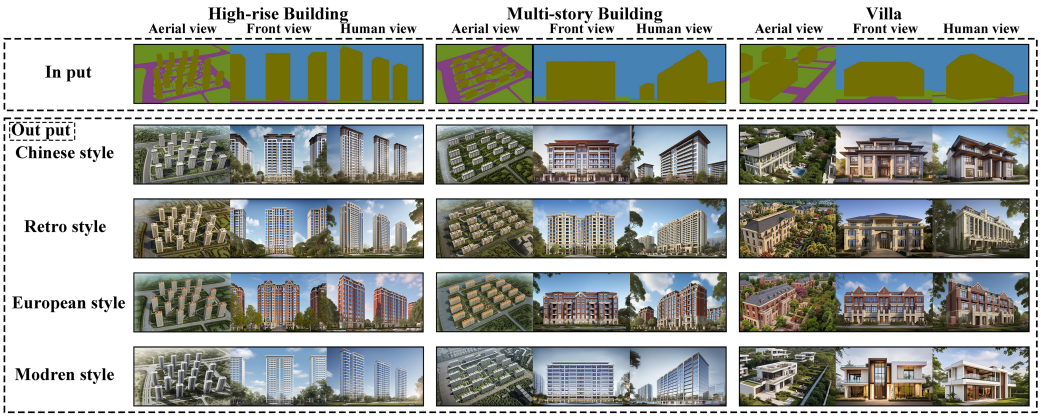

Figure 8. Renderings of residential areas with different building types, styles, and viewpoints generated by GRDPT.

4.3. Evaluation Indicators

Evaluating generated design images is challenging because automated measures of image aesthetics or composition do not fully capture design content [8,14,61,62]. Architectural design encompasses vast expertise and heavily relies on human assessment. Therefore, the seven evaluation criteria in this study were developed from two sources. First, discussions with experienced designers and members of the public informed the selection of criteria by incorporating professional design requirements and viewers' perceptions. Second, our previous studies on AI-generated interior design provided a methodological basis for evaluating usability, style consistency, visual quality, and design plausibility [13,14]. Combining these perspectives, we adapted the previous evaluation frameworks to residential area renderings and established seven criteria covering overall impressions, architectural characteristics, surrounding landscape, and realism.

The established evaluation metrics include usability, building style consistency, first impression, building integrity, building landscape, building details, and realism. The evaluation employs a binary scoring system, where professional designers assess each indicator in the generated images based on their design experience. Images that satisfy the required standards receive a score of 1, while those that fail to meet these criteria are assigned a score of 0. This binary approach facilitates rapid and efficient judgment by designers. Compared with multi-point rating scales such as Likert scales, binary judgments are simpler to make and reduce the differences in how individual evaluators interpret and use the intermediate points of a scale. Previous studies have shown that binary questions are answered more quickly and perceived as less difficult than multi-category answer formats, while yielding equally reliable results [63]. Binary judgments are also widely used in the evaluation of text-to-image generation; for example, the Davidsonian Scene Graph framework evaluates generated images through a set of binary yes/no questions [64]. In addition, the binary scheme corresponds to the practical decision that designers make when screening concept renderings, namely whether a rendering can be used or not.

Next, the content and importance of each evaluation criterion are elaborated. Usability and building style consistency are considered the most critical criteria. The former refers to the observer's satisfaction with all indicators and details of the image for usage in design projects. Building style consistency evaluates the consistency between the generated design image and the architectural style prompt. These two criteria largely determine the overall quality of the generated design image. The first impression assesses the user's overall feeling upon first seeing the image without over-focusing on details, as long as the image appears free of obvious errors.

Building integrity determines whether there are any defects in the building. Building landscape evaluates the completeness and quality of the generated elements other than the building. Building details verify the clarity and rationality of the architectural design details in the image. The realism criterion assesses the similarity between the generated image and a real photo taken by a camera.

4.4. Visual Assessment

This section will visually compare mainstream image generation models, such as Midjourney [51], DALL E3 [52], and Stable Diffusion 1.5 [50], with the GRDPT proposed in this study. Midjourney and DALL E3 are prominent commercial platforms boasting millions of users. Meanwhile, Stable Diffusion has garnered tens of millions of users and offers superior control over the image generation process, making it an essential benchmark for comparison. By comparing these models, we aim to evaluate the effectiveness of our proposed method across multiple dimensions. In order to ensure relative fairness, all were provided with identical reference or control images during the comparison process. The generated visualization of the residential area is presented in Figure 9.



Figure 9. Comparison of residential area images generated by three mainstream image generation models and GRDPT. (Prompt: Chinese-style residential building, multi-story building, human perspective, best quality.)

The settings of the compared models are as follows. Stable Diffusion 1.5 was combined with ControlNet [29] using the segmentation model `control_v11p_sd15_seg`, which is trained on color-coded semantic segmentation maps, and it received the same segmentation images as GRDPT. The segmentation images exported from Grasshopper were input directly without a preprocessor, and the default ControlNet settings were kept: a control weight of 1.0, a starting control step of 0, an ending control step of 1, and the "Balanced" control mode. The same ControlNet settings were used for GRDPT. Because it shares the same base model and the same conditioning information with GRDPT and differs only in the fine-tuning on RAPD-36, it serves as a controllable-generation baseline. Midjourney version 5.2 and DALL E3, which was accessed through its web interface, received the same

segmentation images as reference images. For Midjourney, the image weight was set to its maximum value with `--iw 2` to preserve the composition of the reference image as much as possible, and the random seeds were specified through the `--seed` parameter; the random seed of DALL E3 cannot be specified by users. For each category, all models used the same text prompt, such as the prompt shown in Figure 9. The prompts followed the template "[style]-style residential building, [building type], [viewpoint], best quality". Stable Diffusion 1.5 and GRDPT were run locally in the AUTOMATIC1111 Stable Diffusion WebUI with the DPM++ 2M Karras sampler, 50 sampling steps, the default CFG scale of 7, and fixed random seeds. For each of the 36 categories, the random seed was set to 1000 for the first image and increased by 1 for each subsequent image, so that the 20 images of a category were generated with seeds 1000 to 1019. The same seeds were used for Stable Diffusion 1.5, GRDPT, and Midjourney. Because Midjourney and DALL E3 generate images at a resolution of 1024 $\times$ 1024 pixels, the images generated by Stable Diffusion 1.5 and GRDPT at 512 $\times$ 512 pixels were upscaled to 1024 $\times$ 1024 pixels using high-resolution fix before comparison. The default settings of high-resolution fix were kept, namely the "Latent" upscaler, an upscaling factor of 2, a denoising strength of 0.7, and hires steps set to 0, which means that the second pass used the same 50 sampling steps as the first pass.

Figure 9 shows that Stable Diffusion performs well in controlling the composition of the generated results based on the segmentation map. Still, it lacks an understanding of architectural styles, and the generated buildings have defects and structural errors. Midjourney can generate structurally reasonable and detailed buildings. Yet, it cannot achieve pixel-level control on the generated results based on the semantic segmentation map. As a result, the compositions do not match the original input, and the generated architectural styles are inconsistent with the prompts, deviating from actual application needs. DALL E3 also struggles to precisely control the composition of the generated images. More importantly, it does not understand architectural styles and cannot create designs in the specified style.

In summary, the existing methods still need improvement in terms of architectural style, building integrity, background details, and realism. In contrast, in the examples shown, the residential area renderings generated by GRDPT follow the input composition and the specified architectural style more closely than those of the above methods.

4.5. Quantitative Evaluation

This section quantitatively evaluates the designs generated by the four image generation models, including Midjourney, DALL E3, Stable Diffusion, and GRDPT, using the settings described in Section 4.4. Each of the four models generates images in four architectural styles, three building types, and three architectural perspectives, totaling 36 types ($4 \times 3 \times 3$). In each type, 20 images are produced with a resolution of 1024 $\times$ 1024, giving 720 images per model and 2880 residential area design images in total (4×720) for evaluation.

To reduce the influence of individual preferences on the evaluation, this study invited 10 designers with relevant experience. Each designer classified the generated 2880 images into two categories based on their design experience, using the seven criteria proposed in this study. The 2880 images were presented to each designer in a random order, and the designers were not informed of the model that generated each image. For each criterion, an image received a score of 1 if it met the corresponding requirement and a score of 0 otherwise. An image was regarded as usable if it could be used directly as a concept rendering in a preliminary design presentation without major editing and if its building massing and layout matched the input segmentation image. Building style consistency was satisfied when the dominant architectural style of the image matched the style specified in

the prompt, following the labeling rules described in Section 3.2. A positive first impression was given when no obvious errors could be noticed while viewing the image as a whole for a few seconds. Building integrity required the buildings to be complete, without missing, fused, floating, or structurally implausible parts. Building landscape required the elements other than the buildings, such as greenery, roads, water, and people, to be complete and placed at reasonable positions and scales. Building details required facade elements such as windows, balconies, railings, and materials to appear clear and plausible at normal viewing size. Finally, an image met the realism criterion if it could plausibly be mistaken for a photograph taken by a camera. For each model, the scores that all ten designers gave to its 720 images were summed for each criterion and divided by the total of 7200 ratings, and the result was converted into a percentage. The specific scores of the four image generation models are shown in Figure 10.

Because each image was rated by all ten designers, the ratings were not independent. For each designer, model, and criterion, the binary ratings of 720 images were averaged to obtain a pass rate. GRDPT was compared with each of the three baseline models on seven criteria using two-sided paired t tests, giving 21 comparisons. Pass rates were paired within the same designer, with designers serving as the analysis units ($n = 10$, $df = 9$), rather than treating the 7200 ratings per model and criterion as independent observations. Mean paired differences were reported in percentage points with unadjusted 95% confidence intervals based on the t distribution with nine degrees of freedom. Holm correction was applied jointly to all 21 tests, with adjusted $p < 0.05$ indicating statistical significance. Inter-rater agreement was assessed for each criterion using Fleiss' κ [65]. The κ values ranged from 0.461 to 0.712, indicating moderate to substantial agreement [66], while observed agreement ranged from 73.4% to 86.2%. These results support the consistency of the designers' judgments but do not establish the scores as objective measurements of design quality.

Figure 10 presents the scores, while Table 3 reports the paired comparisons between GRDPT and the three baseline models. In this evaluation, GRDPT obtained higher mean scores in 20 of the 21 comparisons, with 14 differences remaining statistically significant after Holm correction. For usability, GRDPT exceeded Midjourney, DALL·E 3, and Stable Diffusion by 68.31, 65.15, and 33.13 percentage points, respectively (all Holm-adjusted $p < 0.001$). For building style consistency, the corresponding differences were 22.89, 32.43, and 27.00 percentage points (all Holm-adjusted $p < 0.001$). GRDPT also obtained significantly higher realism scores than all three baselines, with differences of 15.00, 9.32, and 41.46 percentage points, respectively (Holm-adjusted $p < 0.01$, $p < 0.001$, and $p < 0.001$).

The results for the remaining criteria were less consistent. For first impression, GRDPT scored significantly higher than Stable Diffusion by 28.31 percentage points (Holm-adjusted $p < 0.05$), whereas its differences from Midjourney and DALL·E 3 were not statistically significant. For building integrity, GRDPT exceeded DALL·E 3 and Stable Diffusion by 28.07 and 40.56 percentage points, respectively (both Holm-adjusted $p < 0.001$), but its difference from Midjourney was not significant. For building landscape, only the comparison with Stable Diffusion was significant, with a difference of 16.76 percentage points (Holm-adjusted $p < 0.01$). For building details, GRDPT exceeded Stable Diffusion by 25.78 percentage points (Holm-adjusted $p < 0.01$). Its difference from DALL·E 3 was not significant, and its mean score was 8.18 percentage points lower than Midjourney, although this difference was also not significant.

Table 3. Comparisons of GRDPT with three baseline models.

Evaluation Criterion	Baseline Model	Mean Difference (pp)	95% CI (Unadjusted)	Holm-Adjusted p Value
Usability	Midjourney	68.31	[61.33, 75.28]	$p < 0.001$
	DALL·E 3	65.15	[58.22, 72.08]	$p < 0.001$
	Stable Diffusion	33.13	[23.37, 42.88]	$p < 0.001$
First impression	Midjourney	4.46	[−8.45, 17.37]	$p \geq 0.05$
	DALL·E 3	9.51	[−2.21, 21.24]	$p \geq 0.05$
	Stable Diffusion	28.31	[10.67, 45.94]	$p < 0.05$
Building style consistency	Midjourney	22.89	[17.07, 28.71]	$p < 0.001$
	DALL·E 3	32.43	[26.58, 38.28]	$p < 0.001$
	Stable Diffusion	27.00	[18.18, 35.82]	$p < 0.001$
Building integrity	Midjourney	5.44	[−4.58, 15.47]	$p \geq 0.05$
	DALL·E 3	28.07	[19.02, 37.12]	$p < 0.001$
	Stable Diffusion	40.56	[33.03, 48.08]	$p < 0.001$
Building landscape	Midjourney	3.57	[−5.85, 12.99]	$p \geq 0.05$
	DALL·E 3	11.47	[−3.48, 26.43]	$p \geq 0.05$
	Stable Diffusion	16.76	[10.14, 23.38]	$p < 0.01$
Building details	Midjourney	−8.18	[−19.37, 3.01]	$p \geq 0.05$
	DALL·E 3	2.58	[−15.10, 20.27]	$p \geq 0.05$
	Stable Diffusion	25.78	[15.09, 36.47]	$p < 0.01$
Realism	Midjourney	15.00	[8.95, 21.05]	$p < 0.01$
	DALL·E 3	9.32	[6.20, 12.44]	$p < 0.001$
	Stable Diffusion	41.46	[31.57, 51.34]	$p < 0.001$

Overall, 14 of the 21 comparisons in this evaluation were statistically significant after Holm correction. GRDPT scored significantly higher than all three baseline models in usability, building style consistency, and realism, showing a consistent advantage across baselines in these three dimensions. Results for the remaining criteria varied by baseline, indicating that the advantage did not extend to every evaluation dimension. Nonsignificant results may reflect subjective judgments, variation among raters, smaller score differences, or the limited number of raters. Therefore, when interpreting these results, inter-rater reliability and confidence intervals should be taken into account; furthermore, it is important to note that a lack of significant difference does not equate to equivalent performance.

4.6. Generated Residential Area Details Demonstration

Figure 11 shows the details and effects of the modern multi-story residential area design generated by GRDPT. The generated building shows symmetry and structural integrity. The generated architectural styles, types, and perspectives are consistent with the input control prompts. A careful observation of the generated design details reveals that the buildings and their accompanying scenery are abundant in quantity, appropriately located, suitably sized, and free from apparent defects. The generated images also exhibit high realism, with realistic light-shadow relationships and building reflections off the water surface. This example illustrates GRDPT's ability to generate high-quality images. However, there is still room for improvement in the model, e.g., the generated building walls and some elements are not completely vertical, with some minor flaws in details. Overall, the example suggests that GRDPT is feasible for the preliminary stage of residential area planning and may help improve designers' efficiency in generating design schemes and making design decisions.

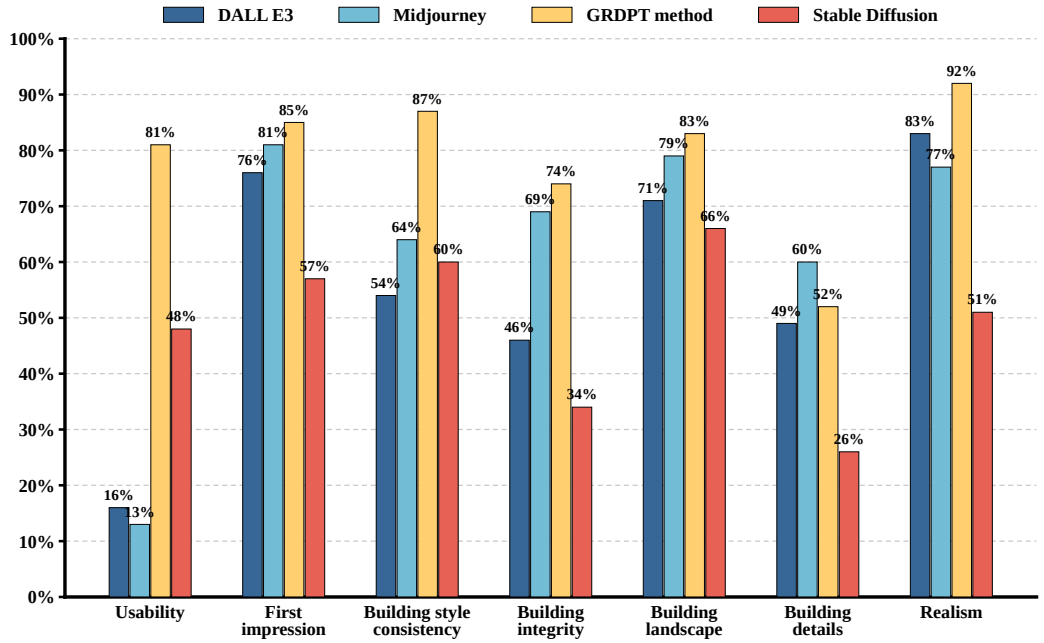

Figure 10. Quantitative evaluation results of residential area renderings generated by four image generation models.



Figure 11. An example of a modern multi-story residential building rendering generated by GRDPT. (Prompt: Modern-style residential building, multi-story building, front view, with a pool in the foreground and greenery around it.)

5. Discussion

This study explores the feasibility of parametric and AI technologies in generating 3D residential area models and renderings and examines GRDPT through a qualitative comparison and a subjective designer evaluation. The qualitative comparison visually compares GRDPT with other mainstream image generation models, illustrating that GRDPT can quickly generate diverse residential area designs with specified building types, styles, and perspectives. In the designer evaluation, GRDPT received higher scores than the compared models in usability, building style consistency, and realism, whereas it received

a lower score than Midjourney in building details. These scores are based on the binary judgments of ten designers. The agreement among the designers supports the reliability of these judgments, but the scores remain subjective and should not be interpreted as objective measurements of design quality. Finally, the details of the residential area plans generated by GRDPT are presented, further illustrating the feasibility of GRDPT.

Overall, the proposed GRDPT provides a practical workflow for residential area planning and design. It transforms the traditional linear workflow into AI-generated designs, reducing the manual work of drawing, modeling, material mapping, and rendering at the preliminary design stage. On a PC with 24 GB of VRAM, GRDPT can generate approximately 6 residential area designs with 3D models and renderings per minute. This suggests that GRDPT can accelerate the production of preliminary design alternatives. Additionally, our method can generate many schemes with different architectural styles, types, and perspectives for designers to choose from, which may support the exploration and comparison of design alternatives.

6. Conclusions

Given the high demands of urbanization on the quality of residential area planning, traditional residential planning and design methods face urgent innovation due to their low efficiency. This study proposes a residential area planning design generation method (GRDPT) based on parametric design and artificial intelligence diffusion models. GRDPT combines the Grasshopper program written by designers with a fine-tuned MARPR model, efficiently generating 3D residential area models and renderings in various architectural styles on a PC. In the subjective evaluation of 2880 generated images by ten designers, GRDPT obtained higher scores than Midjourney, DALL E3, and Stable Diffusion in usability (81%), building style consistency (87%), and realism (92%), whereas it scored lower than Midjourney in building details (52% vs. 60%). With the aid of GRDPT, designers can rapidly generate 3D models and renderings of residential areas at the preliminary design stage. This may improve designers' efficiency in refining design proposals and reduce tedious drawing and modification tasks. This research provides new workflows and tools for generating and modifying residential area plans.

However, this study also has some limitations. Developing comprehensive quantitative evaluation standards for highly subjective design types remains challenging. Although this study initially established quantitative evaluation standards based on relevant literature and expert opinions, future research should broaden the evaluation dimensions to more effectively capture and quantify subjective perceptions. The human evaluation in this study is subjective. Fleiss' κ values of 0.461–0.712 indicate moderate to substantial agreement among the ten designers, supporting rating reliability without establishing objective design quality. Of the 21 comparisons, 14 were statistically significant after Holm correction, with GRDPT scoring significantly higher than all three baselines in usability, building style consistency, and realism. Results for the other criteria varied by baseline, and the lower building-detail score relative to Midjourney was not statistically significant. Nonsignificant differences establish neither superiority nor equivalence. The binary scoring scheme cannot capture differences in degree, and shared evaluator preferences may influence the ratings. Without objective image-quality metrics or real-project performance measures, the findings remain limited to the judgments of the ten designers, the 36 categories, and the model versions tested. Furthermore, the current classification of architectural styles is based on traditional Chinese notions, and its applicability on a global scale still requires verification, leading to the need for a universal classification standard for architectural styles. In practice, generating complex scenes using diffusion models is still a challenge. First, the limited resolution of images generated by diffusion models

constrains their ability to accurately represent intricate scenes. Second, fine-tuning diffusion models or pre-programmed Grasshopper scripts necessitates specialized programming and computational skills, thereby limiting the accessibility of the technology. The method proposed in this study is particularly well-suited for the preliminary design stage, as it enables rapid generation of initial design proposals for discussion. This method enables a rapid determination of the architectural planning layout and overall style, which can subsequently be refined into construction drawings by designers.

Future research could expand architectural style coverage, develop higher-resolution generation models, and improve integration between Grasshopper and diffusion models. Large language models such as GPT-6 Astra [67] could be explored for generating Grasshopper programs from natural-language requirements, reducing programming demands on designers. Integration with image generation models such as GPT Image 2 [68] could further streamline modeling and visualization. Additional architectural design modules and more comprehensive evaluation criteria could support broader application of the workflow. Future studies should strengthen the evaluation by addressing evaluator preferences, the limited panel size, and the coarse resolution of binary ratings. Recruiting architects and planners from different institutions would broaden the perspectives represented, while clearly defined multi-point scales and shared scoring examples could capture differences in degree and improve consistency. Designer ratings should also be complemented by measures of layout adherence based on segmentation overlap, image quality relative to real residential area photographs, and text–image alignment for style consistency. These measures would provide additional evidence rather than fully represent design quality. To assess practical usability, task-based studies should record the time and editing required to develop generated renderings into deliverable design schemes and examine whether the resulting layouts satisfy project requirements. Together, these approaches would connect subjective judgments with measurable visual and practical outcomes.

Author Contributions: Conceptualization, K.Z. and J.C.; methodology, K.Z.; software, J.C.; validation, X.Z. and J.C.; formal analysis, J.C.; investigation, J.C.; resources, J.C.; data curation, K.Z. and J.C.; writing—original draft, K.Z., X.Z. and J.C.; writing—review and editing, K.Z., X.Z. and J.C.; visualization, J.C.; supervision, K.Z. and J.C.; project administration, J.C.; funding acquisition, J.C. All authors have read and agreed to the published version of the manuscript.

Funding: This research work is funded by the FRG-25-064-FA and granted by the Research Fund of Macao University of Science and Technology (FRG-MUST). This work was supported by the Science and Technology Development Fund of Macau Project 0087/2025/ITP2.

Data Availability Statement: The annotations of the RAPD-36 dataset presented in this study are available on request from the corresponding author for academic research purposes. The original images are not publicly available because they were collected from public web sources, and their copyright belongs to their original owners, which restricts their redistribution.

Acknowledgments: During the preparation of this manuscript, the authors used Midjourney 5.2, DALL·E 3 and Stable Diffusion 1.5 in order to visually compare the performance of different AI image generation tools, as presented in Figures 8 and 9.

Conflicts of Interest: The authors declare no conflicts of interest.

References

1. Yang, Z.; Cao, Y.; Du, J. The Impact of New Urbanization Construction on Sustainable Economic Growth of Resource-Based Cities. *Environ. Sci. Pollut. Res.* **2023**, *30*, 96860–96874. [CrossRef]
2. Wu, X.; Huang, Y.; Gao, J. Impact of industrial agglomeration on new-type urbanization: Evidence from Pearl River Delta urban agglomeration of China. *Int. Rev. Econ. Financ.* **2022**, *77*, 312–325. [CrossRef]

3. Li, Q.; Zeng, F.E.; Liu, S.; Yang, M.; Xu, F. The effects of China's sustainable development policy for resource-based cities on local industrial transformation. *Resour. Policy* **2021**, *71*, 101940. [CrossRef]
4. Zhang, J.; Liu, N.; Wang, S. A parametric approach for performance optimization of residential building design in Beijing. In *Proceedings of the Building Simulation*; Springer: Berlin/Heidelberg, Germany, 2020; Volume 13, pp. 223–235. [CrossRef]
5. Sohatskaya, D.G. Parametric architecture and artificial intelligence as a way of modeling three-dimensional environment. In *Socio-Economic Systems: Paradigms for the Future*; Springer: Berlin/Heidelberg, Germany, 2021; pp. 33–39. [CrossRef]
6. Ahmad, A.; Bande, L. AI Application in Architecture in UAE: Optimization of Parametric Structure as a Retrofit Strategy of a Mid-rise Residential Building in Downtown Abu Dhabi. In *Proceedings of the International Conference on Human-Computer Interaction*; Springer: Berlin/Heidelberg, Germany, 2022; pp. 301–313. [CrossRef]
7. Gu, R.; Zhou, W. Application of Parametric Design in Urban Planning. In *Proceedings of the Artificial Intelligence Algorithms and Applications: 11th International Symposium, ISICA 2019, Guangzhou, China, 16–17 November 2019, Revised Selected Papers 11*; Springer: Berlin/Heidelberg, Germany, 2020; pp. 526–540. [CrossRef]
8. Chen, J.; Wang, D.; Shao, Z.; Zhang, X.; Ruan, M.; Li, H.; Li, J. Using artificial intelligence to generate master-quality architectural designs from text descriptions. *Buildings* **2023**, *13*, 2285. [CrossRef]
9. Xia, B.; Li, X.; Shi, H.; Chen, S.; Chen, J. Style classification and prediction of residential buildings based on machine learning. *J. Asian Archit. Build. Eng.* **2020**, *19*, 714–730. [CrossRef]
10. Chen, J.; Shao, Z.; Zhu, H.; Chen, Y.; Li, Y.; Zeng, Z.; Yang, Y.; Wu, J.; Hu, B. Sustainable interior design: A new approach to intelligent design and automated manufacturing based on Grasshopper. *Comput. Ind. Eng.* **2023**, *183*, 109509. [CrossRef]
11. Samuelson, H.; Claussnitzer, S.; Goyal, A.; Chen, Y.; Romo-Castillo, A. Parametric energy simulation in early design: High-rise residential buildings in urban contexts. *Build. Environ.* **2016**, *101*, 19–31. [CrossRef]
12. Tian, J.; Ooka, R. Evaluation of solar energy potential for residential buildings in urban environments based on a parametric approach. *Sustain. Cities Soc.* **2024**, *106*, 105350. [CrossRef]
13. Chen, J.; Shao, Z.; Hu, B. Generating Interior Design from Text: A New Diffusion Model-Based Method for Efficient Creative Design. *Buildings* **2023**, *13*, 1861. [CrossRef]
14. Chen, J.; Shao, Z.; Zheng, X.; Zhang, K.; Yin, J. Integrating aesthetics and efficiency: AI-driven diffusion models for visually pleasing interior design generation. *Sci. Rep.* **2024**, *14*, 3496. [CrossRef]
15. Kim, C.; Cho, S.; Park, Y.; Lee, L.; Lee, J.K. Transforming Architectural Visualizations for Generative xR Renderings using Fine-tuned AI Models. In *Proceedings of the 2023 IEEE MIT Undergraduate Research Technology Conference (URTC)*; IEEE: New York, NY, USA, 2023; pp. 1–5. [CrossRef]
16. Lee, J.K.; Yoo, Y.; Cha, S.H. Generative early architectural visualizations: Incorporating architect's style-trained models. *J. Comput. Des. Eng.* **2024**, *11*, 40–59. [CrossRef]
17. Hamidavi, T.; Abrishami, S.; Ponterosso, P.; Begg, D.; Nanos, N. OSD: A framework for the early stage parametric optimisation of the structural design in BIM-based platform. *Constr. Innov.* **2020**, *20*, 149–169. [CrossRef]
18. Taleb, H.; Musleh, M.A. Applying urban parametric design optimisation processes to a hot climate: Case study of the UAE. *Sustain. Cities Soc.* **2015**, *14*, 236–253. [CrossRef]
19. Wang, S.; Yi, Y.K.; Liu, N. Multi-objective optimization (MOO) for high-rise residential buildings' layout centered on daylight, visual, and outdoor thermal metrics in China. *Build. Environ.* **2021**, *205*, 108263. [CrossRef]
20. Ma, H.; Zheng, H. Text Semantics to Image Generation: A method of building facades design base on Stable Diffusion model. In *Proceedings of the The International Conference on Computational Design and Robotic Fabrication*; Springer: Berlin/Heidelberg, Germany, 2023; pp. 24–34. [CrossRef]
21. Wen, M.; Liang, D.; Ye, H.; Tu, H. Architectural facade design with style and structural features through stable diffusion model. *J. Intell. Constr.* **2024**, *2*, 1–12. [CrossRef]
22. Nichol, A.Q.; Dhariwal, P. Improved denoising diffusion probabilistic models. In *Proceedings of the International Conference on Machine Learning*; Meila, M., Zhang, T., Eds.; Proceedings of Machine Learning Research (PMLR): Cambridge, MA, USA, 2021; Volume 139, pp. 8162–8171. [CrossRef]
23. Choi, J.; Kim, S.; Jeong, Y.; Gwon, Y.; Yoon, S. ILVR: Conditioning Method for Denoising Diffusion Probabilistic Models. In *Proceedings of the 2021 IEEE/CVF International Conference on Computer Vision (ICCV)*; IEEE: Montreal, QC, Canada, 2021; pp. 14347–14356. [CrossRef]
24. Nichol, A.Q.; Dhariwal, P.; Ramesh, A.; Shyam, P.; Mishkin, P.; Mcgrew, B.; Sutskever, I.; Chen, M. GLIDE: Towards Photorealistic Image Generation and Editing with Text-Guided Diffusion Models. In *Proceedings of the International Conference on Machine Learning*; PMLR: Cambridge, MA, USA, 2022; pp. 16784–16804. [CrossRef]
25. Gu, S.; Chen, D.; Bao, J.; Wen, F.; Zhang, B.; Chen, D.; Yuan, L.; Guo, B. Vector quantized diffusion model for text-to-image synthesis. In *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*; IEEE: Montreal, QC, Canada, 2022; pp. 10696–10706. [CrossRef]

26. Ruiz, N.; Li, Y.; Jampani, V.; Pritch, Y.; Rubinstein, M.; Aberman, K. DreamBooth: Fine Tuning Text-to-Image Diffusion Models for Subject-Driven Generation. In *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)*; IEEE: Montreal, QC, Canada, 2023; pp. 22500–22510. [CrossRef]
27. Hu, E.J.; Shen, Y.; Wallis, P.; Allen-Zhu, Z.; Li, Y.; Wang, S.; Wang, L.; Chen, W. LoRA: Low-Rank Adaptation of Large Language Models. *arXiv* **2022**, arXiv:2106.09685.
28. Kumari, N.; Zhang, B.; Zhang, R.; Shechtman, E.; Zhu, J.Y. Multi-Concept Customization of Text-to-Image Diffusion. In *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)*; IEEE: Montreal, QC, Canada, 2023; pp. 1931–1941. [CrossRef]
29. Zhang, L.; Rao, A.; Agrawala, M. Adding Conditional Control to Text-to-Image Diffusion Models. In *Proceedings of the IEEE/CVF International Conference on Computer Vision (ICCV)*; IEEE: Montreal, QC, Canada, 2023; pp. 3836–3847. [CrossRef]
30. Chen, J.; Zheng, X.; Shao, Z.; Ruan, M.; Li, H.; Zheng, D.; Liang, Y. Creative interior design matching the indoor structure generated through diffusion model with an improved control network. *Front. Archit. Res.* **2024**, *14*, 1528. [CrossRef]
31. Shao, Z.; Chen, J.; Zeng, H.; Hu, W.; Xu, Q.; Zhang, Y. A New Approach to Interior Design: Generating Creative Interior Design Videos of Various Design Styles from Indoor Texture-Free 3D Models. *Buildings* **2024**, *14*, 1528. [CrossRef]
32. Addas, A. Role of urban planning standards in improving lifestyle in a sustainable system. *Sustainability* **2023**, *15*, 9171. [CrossRef]
33. Bhattarai, K.; Budd, D. Effects of rapid urbanization on the quality of life. In *Multidimensional Approach to Quality of Life Issues*; Springer: Singapore, 2019; pp. 327–341. [CrossRef]
34. Jiang, Y.; Huang, G. Urban residential quarter green space and life satisfaction. *Urban For. Urban Green.* **2022**, *69*, 127510. [CrossRef]
35. Li, C.; Zhang, T.; Du, X.; Zhang, Y.; Xie, H. Generative AI models for different steps in architectural design: A literature review. *Front. Archit. Res.* **2024**, *14*, 759–783. [CrossRef]
36. Jo, H.; Lee, J.; Lee, Y.; Choo, S. Generative AI and building design: Early photorealistic render visualization of façades using local identity-trained models. *J. Comput. Eng.* **2024**, *11*, 85–105. [CrossRef]
37. Lee, J.K.; Lee, S.; Kim, Y.c.; Kim, S.; Hong, S.W. Augmented virtual reality and 360 spatial visualization for supporting user-engaged design. *J. Comput. Des. Eng.* **2023**, *10*, 1047–1059. [CrossRef]
38. Lee, J.K.; Jeong, H.; Kim, Y.; Cha, S.H. Creating spatial visualizations using fine-tuned interior design style models informed by user preferences. *Adv. Eng. Inform.* **2024**, *62*, 102686. [CrossRef]
39. Murguia, D.; Brioso, X.; Ruiz-Conejo, L.; Fernandez, L. Process integration framework for the design phase of a residential building. *Procedia Eng.* **2017**, *196*, 462–469. [CrossRef]
40. Zhang, J.; Liu, N.; Wang, S. Generative design and performance optimization of residential buildings based on parametric algorithm. *Energy Build.* **2021**, *244*, 111033. [CrossRef]
41. Casini, M. Chapter 6—Advanced digital design tools and methods. In *Construction 4.0*; Casini, M., Ed.; Woodhead Publishing Series in Civil and Structural Engineering; Woodhead Publishing: Cambridge, UK, 2022; pp. 263–334. [CrossRef]
42. Bai, X.; Huerta, O.; Unver, E.; Allen, J.; Clayton, J.E. A parametric product design framework for the development of mass customized head/face (eyewear) products. *Appl. Sci.* **2021**, *11*, 5382. [CrossRef]
43. Brown, N.C.; Jusiega, V.; Mueller, C.T. Implementing data-driven parametric building design with a flexible toolbox. *Autom. Constr.* **2020**, *118*, 103252. [CrossRef]
44. Jeong, J.; Park, H.; Lee, Y.; Kang, J.; Chun, J. Developing parametric design fashion products using 3D printing technology. *Fash. Text.* **2021**, *8*, 22. [CrossRef]
45. Leung, T.M. Parametric design modelling in urban art: Approaches and future directions. In P*roceedings of the 2019 International Conference on Architecture: Heritage, Traditions and Innovations (AHTI 2019)*; Atlantis Press: Dordrecht, The Netherlands, 2019; pp. 457–461. [CrossRef]
46. Zhang, J.; Ji, L. Optimization of daylighting, ventilation, and cooling load performance of apartment in tropical ocean area based on parametric design. *Adv. Civ. Eng.* **2021**, *2021*, 6511290. [CrossRef]
47. Wu, H.; Deng, F.; Tan, H. Research on parametric design method of solar photovoltaic utilization potential of nearly zero-energy high-rise residential building based on genetic algorithm. *J. Clean. Prod.* **2022**, *368*, 133169. [CrossRef]
48. Banihashemi, S.; Tabadkani, A.; Hosseini, M.R. Integration of parametric design into modular coordination: A construction waste reduction workflow. *Autom. Constr.* **2018**, *88*, 1–12. [CrossRef]
49. Brown, N.C.; Mueller, C.T. Design variable analysis and generation for performance-based parametric modeling in architecture. *Int. J. Archit. Comput.* **2019**, *17*, 36–52. [CrossRef]
50. Rombach, R.; Blattmann, A.; Lorenz, D.; Esser, P.; Ommer, B. High-resolution image synthesis with latent diffusion models. In *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*; IEEE: Montreal, QC, Canada, 2022; pp. 10684–10695. [CrossRef]
51. Borji, A. Generated faces in the wild: Quantitative comparison of stable diffusion, midjourney and dall-e 2. *arXiv* **2022**, arXiv:2204.06125. [CrossRef]

52. Betker, J.; Goh, G.; Jing, L.; Brooks, T.; Wang, J.; Li, L.; Ouyang, L.; Zhuang, J.; Lee, J.; Guo, Y.; et al. Improving image generation with better captions. *Comput. Sci.* **2023**, *2*, 8. Available online: https://api.semanticscholar.org/CorpusID:264403242 (accessed on 14 September 2026).
53. Liu, M.; Hu, Y. Application potential of stable diffusion in different stages of industrial design. In *Proceedings of the International Conference on Human-Computer Interaction*; Springer: Berlin/Heidelberg, Germany, 2023; pp. 590–609. [CrossRef]
54. Denton, E.; Hanna, A.; Amironesei, R.; Smart, A.; Nicole, H.; Scheuerman, M.K. Bringing the People Back in: Contesting Benchmark Machine Learning Datasets. *arXiv* **2020**, arXiv:2007.07399. [CrossRef]
55. Pugliese, R.; Regondi, S.; Marini, R. Machine learning-based approach: Global trends, research directions, and regulatory standpoints. *Data Sci. Manag.* **2021**, *4*, 19–29. [CrossRef]
56. Deng, J.; Dong, W.; Socher, R.; Li, L.J.; Li, K.; Fei-Fei, L. Imagenet: A large-scale hierarchical image database. In *Proceedings of the 2009 IEEE Conference on Computer Vision and Pattern Recognition*; IEEE: Montreal, QC, Canada, 2009; pp. 248–255. [CrossRef]
57. Schuhmann, C.; Beaumont, R.; Vencu, R.; Gordon, C.; Wightman, R.; Cherti, M.; Coombes, T.; Katta, A.; Mullis, C.; Wortsman, M.; et al. Laion-5b: An open large-scale dataset for training next generation image-text models. *Adv. Neural Inf. Process. Syst.* **2022**, *35*, 25278–25294. [CrossRef]
58. Delgado, J.M.D.; Oyedele, L.; Ajayi, A.; Akanbi, L.; Akinade, O.; Bilal, M.; Owolabi, H. Robotics and automated systems in construction: Understanding industry-specific challenges for adoption. *J. Build. Eng.* **2019**, *26*, 100868. [CrossRef]
59. Gooood Design Website. Available online: https://www.gooood.cn (accessed on 1 March 2025).
60. ImageNet: Download and Terms of Access. Available online: https://image-net.org/download.php (accessed on 1 March 2025).
61. Wang, W.; Wang, X.; Xue, C. Aesthetics Evaluation Method of Chinese Characters based on Region Segmentation and Pixel Calculation. *Intell. Hum. Syst. Integr. (IHSI 2023) Integr. People Intell. Syst.* **2023**, *69*, 561–568. [CrossRef]
62. Wang, L.; Xue, C. A Simple and Automatic Typesetting Method Based on BM Value of Interface Aesthetics and Genetic Algorithm. In *Proceedings of the Advances in Usability, User Experience, Wearable and Assistive Technology*; Springer: Berlin/Heidelberg, Germany, 2021; pp. 931–938. [CrossRef]
63. Dolnicar, S.; Grün, B.; Leisch, F. Quick, simple and reliable: Forced binary survey questions. *Int. J. Mark. Res.* **2011**, *53*, 231–252. [CrossRef]
64. Cho, J.; Hu, Y.; Garg, R.; Anderson, P.; Krishna, R.; Baldridge, J.; Bansal, M.; Pont-Tuset, J.; Wang, S. Davidsonian Scene Graph: Improving Reliability in Fine-grained Evaluation for Text-to-Image Generation. In *Proceedings of the International Conference on Learning Representations (ICLR)*; ICLR: Appleton, WI, USA, 2024.
65. Fleiss, J.L. Measuring nominal scale agreement among many raters. *Psychol. Bull.* **1971**, *76*, 378–382. [CrossRef]
66. Landis, J.R.; Koch, G.G. The measurement of observer agreement for categorical data. *Biometrics* **1977**, *33*, 159–174. [CrossRef]
67. OpenAI. GPT-6 Astra: A New Generation of Intelligence. Available online: https://openai.com/index/gpt-6-astra/ (accessed on 3 October 2026).
68. OpenAI. Image Generation with GPT Image Models. Available online: https://developers.openai.com/api/docs/guides/image-generation (accessed on 3 October 2026).

Disclaimer/Publisher's Note: The statements, opinions and data contained in all publications are solely those of the individual author(s) and contributor(s) and not of MDPI and/or the editor(s). MDPI and/or the editor(s) disclaim responsibility for any injury to people or property resulting from any ideas, methods, instructions or products referred to in the content.